UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/637,667	08/11/2003	James K. Baker	089348-0160	3520

22428 7590 05/14/2007
FOLEY AND LARDNER LLP
SUITE 500
3000 K STREET NW
WASHINGTON, DC 20007

EXAMINER	
CHAWAN, VIJAY B	
ART UNIT	PAPER NUMBER
2626	

MAIL DATE	DELIVERY MODE
05/14/2007	PAPER

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

Office Action Summary	Application No.	Applicant(s)	
	10/637,667	BAKER, JAMES K.	
	Examiner	Art Unit	
	Vijay B. Chawan	2626	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☐ Responsive to communication(s) filed on _____.

2a)☐ This action is **FINAL**. 2b)☒ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) 1-63 is/are pending in the application.

4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) 1-63 is/are rejected.

7)☐ Claim(s) _____ is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☐ The specification is objected to by the Examiner.

10)☐ The drawing(s) filed on _____ is/are: a)☐ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☐ All b)☐ Some * c)☐ None of:

1.☐ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) ☒ Notice of References Cited (PTO-892)
2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3) ☐ Information Disclosure Statement(s) (PTO/SB/08) Paper No(s)/Mail Date _____.
4) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____.
5) ☐ Notice of Informal Patent Application
6) ☐ Other: _____.

DETAILED ACTION

Claim Rejections - 35 USC § 102

1. The following is a quotation of the appropriate paragraphs of 35 U.S.C. 102 that form the basis for the rejections under this section made in this Office action:

> A person shall be entitled to a patent unless –
>
> (e) the invention was described in (1) an application for patent, published under section 122(b), by another filed in the United States before the invention by the applicant for patent or (2) a patent granted on an application for patent by another filed in the United States before the invention by the applicant for patent, except that an international application filed under the treaty defined in section 351(a) shall have the effects for purposes of this subsection of an application filed in the United States only if the international application designated the United States and was published under Article 21(2) of such treaty in the English language.

2. Claims 1-63 are rejected under 35 U.S.C. 102(e) as being anticipated by Minamino et al., (7,013,277).

As per claim 1, Minamino et al., teach a program product for speech recognition, comprising machine-readable program code for causing, when executed, a machine to perform the following method: detecting at least one speech element in an utterance of acoustic data which meets a criteria for potentially being reduced; presenting at least a portion of the utterance of acoustic data to a user with a prompt to determine if the speech element is reduced; and if verification data is received from the user that the speech element in the utterance is reduced, training an acoustic model for a reduced form of the speech element using data related to the utterance (Col.21, line 39 – Col.22, line 47, Col.19, lines 1-60).

As per claim 2, Minamino et al., teach the program product as defined in claim 1, wherein the criteria is that a portion of the acoustic data of the utterance matches data in a dictionary of reduced speech elements (Col.16, lines 14 – 65).

As per claim 3, Minamino et al., teach the program product as defined in claim 2, wherein the matched data in the dictionary comprises acoustic data for a reduced speech element (Col.16, lines 14 – 65).

As per claim 4, Minamino et al., teach the program product as defined in claim 2, wherein the matched data in the dictionary comprises word sequence context data (Col.16, lines 14 – 65).

As per claim 5, Minamino et al., teach the program product as defined in claim 2, wherein the dictionary is external to a speech recognition model in a base speech recognition process (Col.16, lines 14 – 65).

As per claim 6, Minamino et al., teach the program product as defined in claim 1, wherein said acoustic model for said reduced form is a discriminative model distinguishing said reduced form from the corresponding unreduced form (Col.16, lines 14 – 65).

As per claim 7, Minamino et al., teach the program product as defined in claim 1, wherein the criterion is that the speech element has a substantially lower amplitude than in an unreduced model for that speech element (Col.16, lines 14 – 65).

As per claim 8, Minamino et al., teach the program product as defined in claim 1, wherein the criterion is that the speech element has a substantially shorter duration than an average duration for the speech element (Col.16, lines 14 – 65).

As per claim 9, Minamino et al., teach the program product as defined in claim 1, wherein the criterion is that the speech element has acoustic characteristics that are less extreme than an unreduced model of the speech element (Col.16, lines 14 – 65).

As per claim 10, Minamino et al., teach the program product as defined in claim 1, wherein the criterion is that the speech element has acoustic characteristics that are more like an average speech sound than is an unreduced model for the speech element (Col.16, lines 14 – 65).

As per claim 11, Minamino et al., teach the program product as defined in claim 1, wherein the speech element is a vocalic and the criterion is that the speech element has acoustic characteristics more similar to a uniform acoustic tube than does an unreduced model for the speech element (Col.16, lines 14 – 65).

As per claim 12, Minamino et al., teach the program product as defined in claim 1, wherein the criterion is that the speech element has acoustic characteristics associated with an incomplete articulatory gesture (Col.16, lines 14 – 65).

As per claim 13, Minamino et al., teach the program product as defined in claim 1, wherein the unreduced speech element is formed by closure between the tongue and the roof of the mouth and the criterion is that the speech element has acoustic characteristics associated with only an incomplete or brief contact of the tongue with the roof of the mouth (Col.16, lines 14 – 65).

As per claim 14, Minamino et al., teach a program product for speech recognition, comprising machine-readable program code for causing, when executed, a machine to perform the following method: receiving a training utterance of acoustic data of a word sequence; detecting if the training utterance of acoustic data has at least one speech element which meets a criterion for potentially being reduced; presenting the utterance of acoustic data to a user with a prompt to determine if the speech element is reduced; and if verification data is received from the user that the speech element in the utterance is reduced, then associating a reduced designation with the speech element in the training utterance designated as reduced (Col.21, line 39 – Col.22, line 47, Col.19, lines 1-60).

As per claim 15, Minamino et al., teach the program product as defined in claim 14, wherein the criterion is that the speech element has a substantially lower amplitude than in an unreduced model for that speech element (Col.16, lines 14 – 65).

As per claim 16, Minamino et al., teach the program product as defined in claim 14, wherein the criterion is that the speech element has a substantially shorter duration than an average duration for speech element (Col.16, lines 14 – 65).

As per claim 17, Minamino et al., teach the program product as defined in claim 14, wherein the criterion is that the speech element has acoustic characteristics that are less extreme than an unreduced model for the speech element (Col.16, lines 14 – 65).

As per claim 18, Minamino et al., teach the program product as defined in claim 14, wherein the criterion is that the speech element has acoustic characteristics that are

more like an average speech sound than is an unreduced model for the speech element (Col.16, lines 14 – 65).

As per claim 19, Minamino et al., teach the program product as defined in claim 14, wherein the speech element is a vocalic and the criterion is that the speech element has acoustic characteristics more similar to a uniform acoustic tube than does an unreduced model for the speech element (Col.16, lines 14 – 65).

As per claim 20, Minamino et al., teach the program product as defined in claim 14, wherein the criterion is that the speech element has acoustic characteristics associated with an incomplete articulatory gesture (Col.16, lines 14 – 65).

As per claim 21, Minamino et al., teach the program product as defined in claim 14, wherein the unreduced speech element is formed by closure between the tongue and the roof of the mouth and the criterion is that the speech element has acoustic characteristics associated with only an incomplete or brief contact of the tongue with the roof of the mouth (Col.16, lines 14 – 65).

As per claim 22, Minamino et al., teach a speech recognition process, comprising: detecting at least one speech element in an utterance of acoustic data which meets a criteria for potentially being reduced; presenting at least a portion of the utterance of acoustic data to a user with a prompt to determine if the speech element is reduced; and if verification data is received from the user that the speech element in the utterance is reduced, training a discrimination model using data related to the utterance (Col.21, line 39 – Col.22, line 47, Col.19, lines 1-60).

As per claim 23, Minamino et al., teach the method as defined in claim 22, wherein the criteria is that a portion of the acoustic data of the utterance matches data in a dictionary of reduced speech elements (Col.16, lines 14 – 65).

As per claim 24, Minamino et al., teach the method as defined in claim 23, wherein the matched data in the dictionary comprises acoustic data for a reduced speech element (Col.16, lines 14 – 65).

As per claim 25, Minamino et al., teach the method as defined in claim 23, wherein the matched data in the dictionary comprises word sequence context data.

As per claim 26, Minamino et al., teach the method as defined in claim 22, wherein the dictionary is external to a speech recognition model in a base speech recognition process (Col.16, lines 14 – 65).

As per claim 27, Minamino et al., teach the method as defined in claim 22, wherein said acoustic model for said reduced form is a discriminative model distinguishing said reduced form from the corresponding unreduced form (Col.16, lines 14 – 65).

As per claim 28, Minamino et al., teach the method as defined in claim 22, wherein the criterion is that the speech element has a substantially lower amplitude than in an unreduced model for that speech element (Col.16, line 14 – Col.18, line 65).

As per claim 29, Minamino et al., teach the method as defined in claim 22, wherein the criterion is that the speech element has a substantially shorter duration than an average duration for the speech element (Col.16, line 14 – Col.18, line 65).

As per claim 30, Minamino et al., teach the method as defined in claim 22, wherein the criterion is that the speech element has acoustic characteristics that are less extreme than an unreduced model of the speech element (Col.16, line 14 – Col.18, line 65). As per claim 31, Minamino et al., teach the method as defined in claim 22, wherein the criterion is that the speech element has acoustic characteristics that are more like an average speech sound than is an unreduced model for the speech element (Col.16, line 14 – Col.18, line 65).

As per claim 32, Minamino et al., teach the method as defined in claim 22, wherein the speech element is a vocalic and the criterion is that the speech element has acoustic characteristics more similar to a uniform acoustic tube than does an unreduced model for the speech element (Col.16, line 14 – Col.18, line 65).

As per claim 33, Minamino et al., teach the method as defined in claim 22, wherein the criterion is that the speech element has acoustic characteristics associated with an incomplete articulatory gesture (Col.16, line 14 – Col.18, line 65).

As per claim 34, Minamino et al., teach the method as defined in claim 22, wherein the unreduced speech element is formed by closure between the tongue and the roof of the mouth and the criterion is that the speech element has acoustic characteristics associated with only an incomplete or brief contact of the tongue with the roof of the mouth (Col.16, line 14 – Col.18, line 65).

As per claim 35, Minamino et al., teach a speech recognition method, comprising: receiving a training utterance of acoustic data of a word sequence; detecting if the training utterance of acoustic data has at least one speech element

which meets a criterion for potentially being reduced; presenting the utterance of acoustic data to a user with a prompt to determine if the speech element is reduced; and if verification data is received from the user that the speech element in the utterance is reduced, then associating a reduced designation with the speech element in the training utterance designated as reduced (Col.21, line 39 – Col.22, line 47, Col.19, lines 1-60).

As per claim 36, Minamino et al., teach the method as defined in claim 35, wherein the criterion is that the speech element has a substantially lower amplitude than in an unreduced model for that speech element (Col.16, line 14 – Col.18, line 65).

As per claim 37, Minamino et al., teach the method as defined in claim 35, wherein the criterion is that the speech element has a substantially shorter duration than an average duration for speech element (Col.16, line 14 – Col.18, line 65).

As per claim 38, Minamino et al., teach the method as defined in claim 35, wherein the criterion is that the speech element has acoustic characteristics that are less extreme than an unreduced model for the speech element (Col.16, line 14 – Col.18, line 65).

As per claim 39, Minamino et al., teach the method as defined in claim 35, wherein the criterion is that the speech element has acoustic characteristics that are more like an average speech sound than is an unreduced model for the speech element (Col.16, line 14 – Col.18, line 65).

As per claim 40, Minamino et al., teach the method as defined in claim 35, wherein the speech element is a vocalic and the criterion is that the speech element has

acoustic characteristics more similar to a uniform acoustic tube than does an unreduced model for the speech element (Col.16, line 14 – Col.18, line 65).

As per claim 41, Minamino et al., teach the method as defined in claim 35, wherein the criterion is that the speech element has acoustic characteristics associated with an incomplete articulatory gesture (Col.16, line 14 – Col.18, line 65).

As per claim 42, Minamino et al., teach the method as defined in claim 35, wherein the unreduced speech element is formed by closure between the tongue and the roof of the mouth and the criterion is that the speech element has acoustic characteristics associated with only an incomplete or brief contact of the tongue with the roof of the mouth (Col.16, line 14 – Col.18, line 65).

As per claim 43, Minamino et al., teach a speech recognition system, comprising: a detector for receiving at least one speech element in an utterance of acoustic data which meets a criteria for potentially being reduced; a presentation device for presenting at least a portion of the utterance of acoustic data to a user with a prompt to determine if the speech element is reduced; and a computer training device that, if verification data is received from the user that the speech element in the utterance is reduced, trains an acoustic model for a reduced form of the speech element using data related to the utterance (Col.11, line 63 – Col.13, line 53).

As per claim 44, Minamino et al., teach the system as defined in claim 43, wherein the criteria is that a portion of the acoustic data of the utterance matches data in a dictionary of reduced speech elements (Col.16, line 14 – Col.18, line 65).

As per claim 45, Minamino et al., teach the system as defined in claim 44, wherein the matched data in the dictionary comprises acoustic data for a reduced speech element (Col.16, line 14 – Col.18, line 65).

As per claim 46, Minamino et al., teach the system as defined in claim 44, wherein the matched data in the dictionary comprises word sequence context data.

As per claim 47, Minamino et al., teach the system as defined in claim 44, wherein the dictionary is external to a speech recognition model in a base speech recognition process (Col.16, line 14 – Col.18, line 65).

As per claim 48, Minamino et al., teach the system as defined in claim 43, wherein said acoustic model for said reduced form is a discriminative model distinguishing said reduced form from the corresponding unreduced form (Col.16, line 14 – Col.18, line 65).

As per claim 49, Minamino et al., teach the system as defined in claim 43, wherein the criterion is that the speech element has a substantially lower amplitude than in an unreduced model for that speech element (Col.16, line 14 – Col.18, line 65).

As per claim 50, Minamino et al., teach the system as defined in claim 43, wherein the criterion is that the speech element has a substantially shorter duration than an average duration for the speech element (Col.16, line 14 – Col.18, line 65).

As per claim 51, Minamino et al., teach the system as defined in claim 43, wherein the criterion is that the speech element has acoustic characteristics that are less extreme than an unreduced model of the speech element (Col.16, line 14 – Col.18, line 65).

As per claim 52, Minamino et al., teach the system as defined in claim 43, wherein the criterion is that the speech element has acoustic characteristics that are more like an average speech sound than is an unreduced model for the speech element (Col.16, line 14 – Col.18, line 65).

As per claim 53, Minamino et al., teach the system as defined in claim 43, wherein the speech element is a vocalic and the criterion is that the speech element has acoustic characteristics more similar to a uniform acoustic tube than does an unreduced model for the speech element (Col.16, line 14 – Col.18, line 65).

As per claim 54, Minamino et al., teach the system as defined in claim 43, wherein the criterion is that the speech element has acoustic characteristics associated with an incomplete articulatory gesture (Col.16, line 14 – Col.18, line 65).

As per claim 55, Minamino et al., teach the system as defined in claim 43, wherein the unreduced speech element is formed by closure between the tongue and the roof of the mouth and the criterion is that the speech element has acoustic characteristics associated with only an incomplete or brief contact of the tongue with the roof of the mouth (Col.16, line 14 – Col.18, line 65).

As per claim 56, Minamino et al., teach a speech recognition system, comprising: a receiver for receiving a training utterance of acoustic data of a word sequence; a detector for detecting if the training utterance of acoustic data has at least one speech element which meets a criterion for potentially being reduced; a presentation device for presenting the utterance of acoustic data to a user with a prompt to determine if the speech element is reduced; and logic for, if verification data is received from the user

that the speech element in the utterance is reduced, then associating a reduced designation with the speech element in the training utterance designated as reduced (Col.11, line 63 – Col.13, line 53).

As per claim 57, Minamino et al., teach the system as defined in claim 56, wherein the criterion is that the speech element has a substantially lower amplitude than in an unreduced model for that speech element (Col.16, line 14 – Col.18, line 65).

As per claim 58, Minamino et al., teach the system as defined in claim 56, wherein the criterion is that the speech element has a substantially shorter duration than an average duration for speech element.

As per claim 59, Minamino et al., teach the system as defined in claim 56, wherein the criterion is that the speech element has acoustic characteristics that are less extreme than an unreduced model for the speech element (Col.16, line 14 – Col.18, line 65).

As per claim 60, Minamino et al., teach the system as defined in claim 56, wherein the criterion is that the speech element has acoustic characteristics that are more like an average speech sound than is an unreduced model for the speech element (Col.16, line 14 – Col.18, line 65).

As per claim 61, Minamino et al., teach the system as defined in claim 56, wherein the speech element is a vocalic and the criterion is that the speech element has acoustic characteristics more similar to a uniform acoustic tube than does an unreduced model for the speech element (Col.16, line 14 – Col.18, line 65).

As per claim 62, Minamino et al., teach the system as defined in claim 56, wherein the criterion is that the speech element has acoustic characteristics associated with an incomplete articulatory gesture (Col.16, line 14 – Col.18, line 65).

As per claim 63, Minamino et al., teach the system as defined in claim 56, wherein the unreduced speech element is formed by closure between the tongue and the roof of the mouth and the criterion is that the speech element has acoustic characteristics associated with only an incomplete or brief contact of the tongue with the roof of the mouth (Col.16, line 14 – Col.18, line 65).

Conclusion

3. The prior art made of record and not relied upon is considered pertinent to applicant's disclosure.

Dahan et al., (6,018,708) teach a method and apparatus for performing speech recognition utilizing a supplementary lexicon orthographies.

DeSimone (5,539,861) teaches speech recognition using bio-signals.

Yamada (6,832,192) teaches a speech synthesizing method and apparatus.

Gould et al., (5,920,837) teach a word recognition system which stores two models for some words and allows selective deletion of one such model.

Bennett et al., (6,633,846) teach a distributed realtime speech recognition system.

Even et al., (7,085,716) teach speech recognition using word-in-phrase command.

Takagi (5,819,223) teaches speech adaptation device suitable for speech recognition device and word spotting device.

Gould et al., (5,915,236) word recognition system which alters code executed as a function of available computational resources.

Bahl et al., (5,884,259) teach a method and apparatus for a time synchronous tree-based search strategy.

Chou et al., (5,797,123) teach a method of key-phrase detection and verification for flexible speech understanding.

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Vijay B. Chawan whose telephone number is (571) 272-7601. The examiner can normally be reached on Monday Through Friday 6:30-3:00.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Richemond Dorvil can be reached on (571) 272-7602. The fax phone number for the organization where this application or proceeding is assigned is (571) 273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free).

Vijay B. Chawan
Primary Examiner
Art Unit 2626

VIJAY CHAWAN
PRIMARY EXAMINER

vbc
5/8/07